

Provided for non-commercial research and educational use only. The published version of this paper is available online via the following link:

<http://link.springer.com/article/10.1007%2Fs10441-014-9233-9>

Hylomorphism and the Metabolic Closure Conception of Life

James DiFrisco

Institute of Philosophy, University of Leuven

Kardinaal Mercierplein 2 - box 3200, 3000 Leuven, Belgium

email: james.difrisco@hiw.kuleuven.be ; tel. +32 488 21 09 10

Abstract: This paper examines three exemplary theories of living organization with respect to their common feature of defining life in terms of metabolic closure: autopoiesis, (M, R) systems, and chemoton theory. Metabolic closure is broadly understood to denote the property of organized chemical systems that each component necessary for the maintenance of the system is produced from within the system itself, except for an input of energy. It is argued that two of the theories considered—autopoiesis and (M, R) systems—participate in a hylomorphist pattern of thinking which separates the “form” of the living system from its “matter.” The analysis and critique of hylomorphism found in the work of the philosopher Gilbert Simondon is then applied to these two theories, and on the basis of this critique it is argued that the chemoton model offers a superior theory of minimal life which overcomes many of the problems associated with the other two. Throughout, the relationship between hylomorphism and the understanding of living things as machines is explored. The paper concludes by considering how hylomorphism as a background ontology for theories of life fundamentally influences the way life is defined.

Keywords: hylomorphism – autopoiesis – (M, R) systems – chemoton – metabolic closure – Simondon – machine

1 Introduction: Life and Metabolic Closure

The question “what is life?” would intuitively seem to be a fundamental research question in the life sciences. In recent history, however, due largely to the successes of the reductionistic program in molecular biology, the question itself fell out of favor. “Life” was considered an outdated metaphysical concept which the new science had rendered unnecessary. As the well-known molecular biologist François Jacob wrote in 1970: “Today we no longer study Life in our laboratories.” A similar sentiment was expressed elsewhere years later: “Today a molecular biologist has no need, so far as his work is concerned, for the word life” (Atlan et al. 1994).¹ More recently, Machery (2012) has argued that the very project of defining life is deeply flawed and should be abandoned. Nevertheless, for a variety of reasons, there is a renewal of interest in this fundamental question. This is partly due to the perceived need for a clearly defined concept of life to guide relatively new research fields within biology, such as artificial life, astrobiology, and synthetic biology.² More generally, it is increasingly being recognized that theoretical biology as a whole is in need of a guiding vision of what life is, and that this is not adequately given by the modern evolutionary synthesis, on the one hand, nor by molecular biology, on the other.³

Whether or not one considers the aim of defining life to be worthwhile, presumably most will agree that work on fundamental theories of life is worthwhile. A fundamental theory of life would be one that attempts to describe the most fundamental properties that distinguish living from non-living systems, such that these properties are capable of grounding and explaining the other properties unique to living systems. In practice, the process of proposing and critiquing definitions of life forms an integral methodological part of this enterprise. This process may be valuable for gaining a better understanding of what life is, even if it does not issue in a widespread consensus on a definition or settle the matter once and for all.

Contemporary theoretical approaches to defining what life is differ widely. Morange (2008: 141) provides a useful classification of approaches in terms of their priority ordering of three fundamental properties of living systems—namely, (1) reproduction or replication, (2) macromolecular structure, and (3) metabolism. This paper examines a set of fundamental theories that prioritize the third of these by defining life in terms of “metabolic closure.” A recent series of articles by Letelier (2003, 2006, 2011), Cornish-Bowden (2007), Cárdenas

¹ Cited in Cornish-Bowden, et al. (2007: 840).

² Cf., Bedau (1996), Cleland and Chyba (2002), Chyba and McDonald (1995), Szostak et al. (2001).

³ Among the very many authors who express this view, cf., e.g., Woese (2004), Morange (2008), Kauffman (1993), Cornish-Bowden, et al. (2007), Maturana and Varela (1980) Rosen (1991), Gánti (2003).

(2010), and others have gathered various theories together under this category.⁴ They include: autopoiesis (Maturana and Varela 1980), (M, R) systems (Rosen 1991), the chemoton (Gánti 2003), the hypercycle (Eigen and Schuster 1977), autocatalytic sets (Kauffman 1993), RAF sets (Hordijk and Steel 2004), and others. Broadly speaking, metabolic closure may be defined as the property of metabolic systems, consisting of the transformation and degradation of chemical components, that each metabolite or enzyme needed for the maintenance of the system is produced by the system itself, apart from an influx of energy. In a system of metabolic closure, an initial source of energy is transformed along a series of chemical reactions that are organized in a circular pattern, the products of one reaction being the reactants for the next, allowing the system to persist in time. A metabolism without the structure provided by closure, by contrast, would remain a type of chemical “dissipative structure” as described by nonequilibrium thermodynamics, whose structure is imposed and maintained by the system’s boundary conditions.⁵ The property of closure in a metabolic system allows it to maintain its own structure and grants it a degree of autonomy and cohesion that is not found in other physical systems. To conceive of this property as the most fundamental feature of life is to argue, in effect, that reproduction or replication (and hence evolution) should be considered the achievement of a system that survives long enough to reproduce itself, or to develop the capacity to do so.

This metabolic closure conception of life also generally makes little to no reference to macromolecular structure, to DNA or RNA, nor to any specific biomolecules, nor for that matter to any determinate “matter.” Instead, it postulates that the essence of life is to be found in the “form” of its organization, which may be multiply realized in many different material substrata. In this respect, it draws on *hylomorphism* as a background ontology.⁶ Hylomorphism in conceptions of life tends to arise from the combined refusal to reduce life to its material substratum and the recognition that life cannot be any type of thing or substance distinct from matter: instead, life is to be found in the form or organization of the material substratum. In the models of metabolic closure that we will consider, abstracting from the matter in favor of the form allows for highly general and compact definitions of life. At the same time, for this reason, the models tend to be formalistic and abstract, and far removed

⁴ The system-theoretic concept of “closure,” used to differentiate levels of organization in nature, has also appeared in the literature with increasing frequency, being the theme of a relatively recent international workshop. Cf. Chandler and Van de Vijver (2000).

⁵ Cf., e.g., the discussion of the Belousov-Zhabotinsky reaction in Prigogine and Nicolis (1977).

⁶ The distinction between information and matter, which has often been used to distinguish living from non-living systems, has rightly been recognized as being involved in a kind a modern hylomorphism (cf. Oyama 2000: 1), but the same also holds for the distinction between organization and matter as it tends to be used in fundamental theories of life.

from what is studied by experimental biologists. This is doubtless one reason why they have had little impact in biology thus far, despite the fact that they offer promising resources for approaching the question of what life is.

This paper aims to critically interpret and compare three theories of metabolic closure—autopoiesis, (M, R) systems, and the chemoton—with respect to their hylomorphic character. There are several motivations behind this project. For one thing, these theories continue to influence more recent fundamental theories emphasizing metabolic closure, in part because they remain some of the deepest and most elaborate theories available.⁷ As for why their hylomorphism should be a criterion in their evaluation, the fact is that the most serious problems with these theories are also problems with hylomorphism more generally. In analyzing them in these terms, then, we will have not only analyzed these particular models, but also will have indicated problems facing *any* efforts to call upon on hylomorphism as a background ontology for biology. The argument will be that autopoiesis and (M, R) systems are involved in a problematic kind of hylomorphism that is largely overcome with chemoton theory. We begin by introducing what hylomorphism means in this context before turning to a comparative analysis of the three theories.

2 What is Hylomorphism? Simondon's Critique

Hylomorphism broadly refers to any doctrine which represents an entity as being divided into a form (*morphē* or *eidos*) and a matter (*hyle*). It originated in Aristotle's metaphysics and formed an integral part of Aristotelian philosophy, to such an extent that any discussion of hylomorphism must make reference to Aristotle. Nevertheless, we wish to abstract hylomorphism from the particular nuances of Aristotelianism in order to consider it as a general tendency of thought. In this perspective, hylomorphism represents an established background ontology with a set of ready-made concepts available to the theorist. This corresponds to the understanding of hylomorphism in the work of the French philosopher Gilbert Simondon (1924-1989).

⁷ The theories of autopoiesis and (M, R) systems remain especially influential in the field of artificial life and computational approaches to minimal life. (Cf., e.g., Etcheberry (2004), Letelier et al. (2003, 2006,) Mossio et al. (2009), Cottam et al. (2007).) They also form some of the theoretical background for recent fundamental theories emphasizing organizational closure and autonomy as fundamental biological properties. (Cf., Mossio et al. (2010), Letelier et al. (2011), Ruiz-Mirazo and Moreno (2012).) Chemoton theory, on the other hand, is the preferred theory of minimal life in Maynard Smith and Szathmáry's influential work on major evolutionary transitions (1995), and is also promoted in Dennett (2011, 2013).

The treatment of hylomorphism in what follows will rely on the critique provided by Simondon, particularly his *L'individu et sa genèse physico-biologique* (1995).⁸ This critique is specifically helpful for the models we will examine as it reveals the precise problematic points to be identified in these models. It is not necessary to be a Simondonian to appreciate the force of his critique, but some background on his philosophy may be illuminating. Simondon, whose work is not well-known in the English-speaking world (in part due to its being mostly untranslated), is notable for having developed a theory of individuation across different levels of organization in nature in the manner of a general systems theory—physical, biological, and in later work, psychic and collective. His work may perhaps be best characterized as a process philosophy in the tradition of Bergson and Whitehead, which attempts to provide a systematic framework describing the generative processes of individuation that make an individual what it is. Simondon's main thesis claims that individuation has traditionally been conceived more or less explicitly according to two deeply rooted tendencies of thought: substantialism and hylomorphism (2005: 23). Substantialism is typified in the ancient Greek atomists, and it represents individuation by the action of pre-given atomic individuals. Hylomorphism, on the other hand, represents individuation as the effect of a pre-determined form acting on a passive matter to produce the individual hylomorphic compound. These tendencies of thought are dialectically related to one another, in the sense that the one is seemingly only overcome by recourse to the other, and this is due to the fact that they hold some hidden postulates in common. In Simondon's argument, the most important of these is that they locate the principle of individuation in some individual—atoms, matter or form—instead of in the processual operation of individuation itself. Accordingly, both are attached to a “static” rather than “genetic” or generative conception of individuation.

The critique of these tendencies retains a degree of generality even when taken outside of the context of the problem of individuation. They may be translated into the terms of philosophical and scientific debates concerning emergence, atomism versus holism, downward or formal causation, levels of organization, etc. Generally speaking, substantialism corresponds to reductionism, physicalism, and mechanism, and hylomorphism to anti-reductionism, emergentism, organizational or relational thinking, and even in some cases, “vitalism.” Hylomorphism evidently admits of much more variation in how it may be constructed. Much of this is due to the variety of ways that “form” can be conceived, and the

⁸ This work, which was Simondon's main doctoral thesis, is contained in a later re-issue entitled *L'individuation à la lumière des notions de forme et d'information* (2005), which also includes later works and which will be cited here.

ways it can be distinguished from the matter. Some of this variety may be preserved in order to keep the notion of hylomorphism general; however, as it is the primary opponent here, a basic conceptual scheme for hylomorphism will be sketched in what follows.

The notion of the individual as a composite of form and matter arises in Aristotle's philosophy largely from two different lines of inquiry: the doctrine of the four causes in the *Physics* (repeated in the *Metaphysics*), and the inquiry into what constitutes the "substance" of a thing in Book Z of the *Metaphysics*.⁹ The difference between these two corresponds to different perspectives from which form and matter may be distinguished.¹⁰ Form can be treated as the universal, the trope, or essence, and matter as the particular, the substratum or subject of predication. This corresponds to the inquiry of the *Metaphysics* and the perspective from which hylomorphism is usually treated in contemporary discussions.¹¹ Alternatively, form and matter may be treated as different kinds of "causes" of natural events, as in the *Physics*. As is well-known, Aristotle's concept of causes (*aitia*) was much broader than the more modern notion of what constitutes a cause, which is often restricted to the efficient cause. The causes for Aristotle, on the contrary, were each answers to "why-questions"—being at once explanatory principles and real features of the things being explained. In this context, matter or the material cause consists in the elements of which something is made (earth, air, fire, water), and is the bearer of its passive potential to be changed. The formal cause is at once the essence or definition of the thing, as well as the "shape" and the set of morphological qualities that differentiate the matter. The sense of hylomorphism within this perspective is what will primarily concern us here. In particular, the notion of form as shape may be readily assimilated to the more contemporary concept of organization, or to the set of relations among the material parts which provides a distinct explanatory principle from those principles governing the "matter" (i.e., those of physics and chemistry).

This perspective is of course not unrelated to questions about universals and particulars, or essence and accident. The way form and matter are configured as to their causal importance fundamentally influences how the essence of the individual is defined. While in the *Metaphysics* the form is the essence by definition (1032b1), the form in the sense of the shape or the organization may also be more "essential" for some types of individuals than for others.

⁹ In a sense, Aristotle's hylomorphism could be considered a type of substantialism since it originates in the inquiry of the *Metaphysics* into what constitutes "substance," the latter being understood in a broad sense to include the essence, the substratum, and the "thisness" *hic et nunc*. However, Simondon's "substantialism" is more narrowly tied to atomism and to the rejection of "form."

¹⁰ Cf., Barnes (2003) on the ambiguity in the concept of "form" between the form as an abstract entity or universal, and the form as a particular or concrete entity.

¹¹ Cf., Rea (2011), Johnston (2006), Tahko (2012).

Such is the case, I argue, primarily for fabricated things and machines, where the form makes the thing what it is and secures its function, and the matter needs only to be compatible with it. In fact, this is what is claimed in the formal conceptions of life we will consider: namely, that the essence of life is to be found in its form, or in its organization or relational topology. It is *this* idea that the form is the essence and the matter is the accident that we will understand by the term “hylomorphism.”¹² Everything else seems to follow from this idea: if the form is what is essential, it should be separately determinable from whatever matter would realize it. The explanation of an individual’s distinctive properties may then also rely primarily on the formal cause. Moreover, if the matter need only be compatible with the form, one would expect the form to be “multiply realizable” in different individuals. These are each aspects we will encounter further below.

Keeping these ideas in mind, we may distinguish between two different ways of modeling the minimal living system: a bottom-up approach starting from the matter, and a top-down approach starting from the form.¹³ In a bottom-up approach, one starts from a definite matter—e.g., for the minimal living system, different biomolecules, metabolites, enzymes, etc.—and attempt to construct a living system by arranging the parts in an appropriate way. This is, roughly speaking, the approach utilized by the chemoton model. Alternatively, in a top-down approach, one starts not from any determinate parts, but rather from the *relations* that may obtain among any parts whatsoever that would produce the living system as a whole. This is the strategy of autopoiesis and especially of (M, R) systems. Each approach has its advantages and disadvantages. The feasibility of a bottom-up approach depends on an advanced state of knowledge of the matter, such that it is known how the determinate material parts are arranged to produce the whole organized system. By contrast, a top-down approach is able to model the system by purely abstract or formal means without having to go through the more difficult task of specifying how it works materially. For this reason, however, it is made to depend on the prospects that (1) a material realization can actually be found for the formal system that is modeled, and (2) the specific features of the matter will not need to enter into the determination of the relations that have defined the system in the first place. In other words, this strategy of modeling the form apart from the matter counts on the very possibility of this separation in the real individual.

¹² Hence, it is not enough that form and matter are distinguished in any way for a theory to be hylomorphic. Since any individual at all (and even abstract objects) can be viewed as a hylomorphic compound, this would include “trivial” instances of form where, for instance, form is an aggregate of parts or is reducible to material and/or efficient causes.

¹³ This differs from Szathmáry’s (2005) distinction between top-down and bottom-up approaches to minimal life, where top-down means starting from existing cells and stripping them down to a viable minimum.

How exactly this is worked out will be discussed below in greater detail once the different models of metabolic closure are presented. First, now that we have a general idea of what hylomorphism means, it is necessary to briefly explain how it might be considered a problematic conceptual scheme. Simondon's critique of hylomorphism in *L'individu et sa genèse physico-biologique* comprises several aspects. There is, first, a general critique of form in the sense of the essence or definition, which secures the connection of hylomorphic conceptions of the individual with a kind of static essentialism. Aristotle's hylomorphism is certainly tied to the notion that the forms to be found in nature are ordered in a fixed hierarchy of types, such that the form pre-exists its materialization.¹⁴ According to a more naturalistic perspective, however, form cannot be present from the start as a fixed type or ready-made design, but is only the end result of a generative process. As Simondon argues, hylomorphism seizes upon only the "extremes" of the operation of individuation—an initial matter and the end result (2005: 46). The generative process making an individual, however, occurs *between* the extremes of matter and form (2005: 48, 55, 60). In Simondon's model, over-simplifying somewhat, individuation takes place when an already informed matter—a matter bearing its own "information" (singular or critical points and micro-level properties)—comes into contact with a metastable source of potential energy which is "transduced" or grafted onto the individual and amplified into some eventual form. His paradigmatic example for this process is that of crystallization in a supersaturated solution. Here, the crystalline germ provides the "information" or the singular point where the potential energy of the solution can be "transduced" to compose progressive layers, amplifying the initial germ into the resulting formed individual crystal (2005: 77 ff.).

The hylomorphic separation of the "extremes" of form and matter has its anthropological origin, in Simondon's argument, in human techniques of fabrication, in which the design of a thing is detached from the material used to construct it (2005: 39 ff.). In the present context of minimal life, this origin is especially relevant to the persistence of the machine analogy for understanding organisms.¹⁵ The basic idea is that living systems are fundamentally like machines in that they are defined by their functional organization, or "form," and not by the matter that realizes it. Unlike ordinary non-living things, then, the behavior of living systems and machines is amenable to purely formal or relational explanation, as opposed to physical or dynamical explanation. In this respect, there is a mutual reinforcement between the machine analogy and the hylomorphic conception of life. The analogy between living things

¹⁴ Cf., e.g., *Metaphysics* 1033b18: "That which is spoken of as form or substance is not produced, but the concrete thing which gets its name from this is produced" (Aristotle 1984).

¹⁵ For a recent systematic criticism of the machine conception of the organism, see Nicholson (2013).

and machines appears in each of the models considered below, though with widely different results, either strengthening or undermining the hylomorphism with which it is bound up.

3.1 Autopoiesis

“Autopoiesis” refers to the theory of living organization pioneered by the Chilean biologists Humberto Maturana and Francisco Varela starting in the 1970’s. The theory was formed from the confluence of several factors—notably, the cybernetic ideas of Heinz von Förster, and the anti-representationalist theory of cognition developed by Maturana. The former offered a certain novel way to describe systems through functional and relational analysis as opposed to physical-dynamical methods. This was then coupled with Maturana’s ideas on cognition, derived from extensive studies of neurophysiology. Maturana argued that the function of the nervous system is not to process information or procure representations, but primarily to produce movements coherent with the internal sensori-motor state of the organism (Letelier et al. 2011: 106). This idea was to spawn a series of radical notions concerning the autonomy and self-referentiality of living and cognitive systems, of which perhaps the most important would be the autopoietic theory of life.

The autopoietic theory of living organization represents an answer to the question “what is life?” and in addressing this question it departs from the frameworks of evolutionary biology and molecular biology. In the introduction to the most well-known exposition of their theory (Maturana and Varela 1980), the theme of autopoiesis is presented along a series of oppositions within the science of life—identity and diversity, autonomy and reproduction, and relations and components. The first two of these oppositions reflects their departure from the framework of neo-Darwinian evolutionary biology. The opposition between components and relations, on the other hand, derives from their distinction between two types of explanation: explanations referring to specific physical entities and their properties, and explanations referring to the *relations* that may obtain between entities to produce a given effect.¹⁶ This distinction closely parallels the “bottom-up” and a “top-down” distinction made above. Whereas explanations in terms of the matter are appropriate to the phenomena of physics, Maturana and Varela argue, explanations in terms of the relations are appropriate to the types

¹⁶ “In our explanation of the organization of living systems, we shall be dealing with the relations which the actual physical components must satisfy to constitute [a living system], not with the identification of these components” (76).

of machines described by cybernetics (their example is that of a control plant).¹⁷ The living system is such a machine: it is an “autopoietic” machine describable by a relational topology rather than by the dynamical laws governing material bodies.

The aim of the autopoietic theory is thus to explain the autonomy and identity of living systems while also defining a criterion by which to distinguish the living from the non-living. This criterion is given in the name “autopoiesis,” a neologism meaning “self-producing.” Autopoiesis, the property that each component of the system is produced through the activity of the system itself, is advanced as a “necessary and sufficient” condition of life (1980: 82). Maturana and Varela define this condition more precisely as follows:

An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces the components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network. (1980: 78-9)

According to the first aspect of this definition, in an autopoietic system there is reciprocal causation between the components of the system and relations between them, or their processes of production. If this were not the case, that is, if a crucial component were not produced by the network, or if it failed to contribute to maintenance of the network that produced it, the autopoietic system would no longer produce “itself” over time, and would perish. Of course, the second law of thermodynamics dictates that any real physical system in which the property of autopoiesis might be found would have to produce—and be produced by—more than just “itself.” But this does not compromise the invariant identity of the system’s autopoiesis. For Maturana and Varela, autopoiesis is not a property of the “structure” of the machine, the actual relations obtaining in its physical space, but of its “organization,” the relations defining it as an entity belonging to a certain class (autopoietic or allopoietic).¹⁸ The tautological character of this definition is readily apparent: autopoiesis is that property of the organization of a system that makes it belong to the class of autopoietic systems. More importantly, however, its clean separation from structure renders autopoietic

¹⁷ The passage is worth quoting in full: “To explain the movement of a falling body one resorts to properties of matter, and to laws that describe the conduct of material bodies according to these properties (kinetic and gravitational laws), while to explain the organization of a control plant one resorts to relations and laws that describe the conduct of relations. In the first case, the elements used in the explanation are bodies and their properties; in the second case, they are relations and their relations, independently of the nature of the bodies that satisfy them” (75-6).

¹⁸ “The organization of a machine (or system) does not specify the properties of the components which realize the machine as a concrete system, it only specifies the relations which these must generate to constitute the machine or system as a unity. Therefore, the organization of a machine is independent of the properties of its components” (Maturana and Varela, 1980: 77). Cf. also Maturana (1980: 48).

organization an abstraction removed from the domain of real interactions and actual components.

The second aspect of the definition, which is increasingly emphasized in later formulations (Varela 1994), stipulates that the autopoietic system is real, as opposed to being a computer simulation, and that it produces a boundary distinguishing itself from its environment. This autopoietic production of a boundary, such as a membrane, lies at the root of the autonomy of living systems, since it protects the autopoietic organization against destructive influences from the environment. Together with the first aspect of the definition, this makes the autopoietic system a closed and self-referential machine, which specifies its own internal states and the domain of its changes (Letelier et al. 2003: 266). This points back to Maturana's ideas on the function of the nervous system: living and cognitive systems do not process inputs from outside of the system, but instead keep the system within a given range of internal states. That is to say, the living system as an autopoietic system is primarily homeostatic, and only secondarily metabolic. Maturana and Varela explicitly compare the autopoietic machine to a class of machines studied by cybernetics—homeostatic machines, whose function is to keep a certain variable constant (such as temperature for a thermostat) (1980: 78). Whereas homeostatic machines have as their object some external parameter, however, the autopoietic machine is a homeostatic machine whose constant variable is its own autopoietic organization as such (1980: 79).

Maturana and Varela infer much more about the nature of life from these initial ideas, much of it concerning the priority of autopoiesis over all other biological properties, but the basic outlines of the theory are visible already. It is clear that the theory of autopoiesis is a hylomorphic theory of life in the sense described above. The concept of autopoiesis is introduced by means of a strong opposition between the relations, or the network of relations, and the underlying components of the system. Moreover, it is the form of the relations alone that defines a machine as being autopoietic or allopoietic. The specification of components, on the other side, is taken to be unnecessary for the living machine to find an adequate explanation (1980: 75). The form is thus separately determinable from whatever matter would realize the autopoietic system. This separation between the form and the matter in the autopoietic theory is here a reflection of a methodological division between different types of explanation—i.e., cybernetic versus dynamical, or relational versus material.

Another indication of its hylomorphism is to be found in the presence of formal causality in the theory. This is where matters become more subtle, however. Many of the central problems surrounding the theory of autopoiesis are rooted in ambiguities concerning how

much the theory is attempting to explain, and how much it is merely stipulating definitions. As we noted, the definition of autopoiesis appeared to invoke a tautology. Nevertheless, the authors seem to believe that their theory constitutes an “explanation of the organization of living systems” (1980: 76) and a “characterization of living systems that explains the generation of all the phenomena proper to them” (1980: 83). This theoretical basis is then used in the argument about the primacy of identity and autonomy over other properties of living systems. Looking for the fundamental explanation, however, one finds only the *definition* of autopoietic systems. For biological phenomena, an explanation which makes no use of thermodynamic, chemical or energetic considerations, and which describes no functional mechanisms, is a strange kind of explanation indeed.

It is, in fact, an explanation by formal causality. Autopoietic systems are taken to function as homeostatic machines, having their own organization as a constant variable.¹⁹ For standard homeostatic machines, what allows the system to maintain a variable within a given range are specific mechanisms which detect the value of the variable and “feed back” the information into a corrective action altering its value.²⁰ This may be understood for a variety of possible variables. But for the “variable” of autopoietic organization, no such mechanism is indicated. The closest one comes to an explanation of this extraordinary operation is the notion that, if the autopoietic system did not do this, it would not be autopoietic by definition.²¹ So one is left with an explanation from formal causality: the autopoietic system operates in this manner because it is in its definition to do so. This same kind of explanation by definition is present in the idea of reciprocal causation between the parts and whole, which Varela claims enables the theory to “eliminate the traditional opposition between the constituent components and the global properties” (2000). Since the components are left indeterminate, however, this remains a nominal solution. It is easy to make the components obey a given rule when they can be defined however one likes. On the contrary, the claim of reciprocal causation depends on the possibility of finding realistic components that can satisfy the relations and make a truth from the formal definition.

This cluster of problems points to a deeper issue. Since the network of relations is specified in advance, while the components are left indeterminate, the two can only enter into an extrinsic relation with one another. For Simondon, this is an important feature common to

¹⁹ “They subordinate all changes to the maintenance of their own organization, independently of how profoundly they may otherwise be transformed in the process” (1980: 80).

²⁰ Cf., Wiener (1961).

²¹ “It is thus clear that the fact that autopoietic systems are homeostatic systems which have their own organization as the variable that they maintain constant, is a necessary consequence of the autopoietic organization.” (1980: 80)

both substantialism and hylomorphism: when some element is fully determined prior to the generative process making the individual, it can bear only a secondary or external relation to the remainder of the individual. Hence, the determinate atoms of substantialism bear an extrinsic relation to their global configuration, and the determinate forms of hylomorphism bear an extrinsic relation to their matter. The terms are related by a mere *rapport* rather than a genuine *relation* (Simondon, 2005: 32). A generative *relation*, on the contrary, is internal to the terms related. In the context of minimal living systems, this means that a determinate material component specifies the types of interaction it may enter into with other determinate material components. If this is the case, there should be an intermediary order of processes between the extremes of the abstract relational form and the atomistic component, one which may be more essential for the explanation of living organization than either extreme alone. This intermediary order does not exist within the autopoietic theory.²²

The nature of this intermediary order will be explored further below. Of course, it is true that autopoiesis is not constructed as a generative theory concerned with the origin of life. It is significant, however, that the exploration of the intermediary order we have pointed out would reveal processes which are not the properties of a self-referential machine and which take place prior to the organizing distinction between self and other. The one-sided emphasis on identity, autonomy, and self-referentiality in the theory of autopoiesis (which has even been called “solipsistic”²³) may thus be an effect of its static rather than generative picture of the living system. Extracting the form from the concrete system as a well-defined unity then allows the system to be compared without hesitation to a machine, since in the fabrication of a machine the form is decided in advance and only needs to be realized by a matter. If our critique is granted, however, then perhaps the machine analogy is fundamentally mistaken. This is one of the important results of Robert Rosen’s work on (M, R) systems.

3.2 (M, R) SYSTEMS

(M, R) systems, or “metabolism-repair” systems, is the theory of living organization developed by the mathematical biologist Robert Rosen, whose work is summarized in the compact and difficult book *Life Itself* (1991). It shares much with the theory of autopoiesis:

²² “The establishment of an autopoietic system cannot be a gradual process; either a system is an autopoietic system or it is not. . . . Accordingly there are not and there cannot be intermediate systems.” (1980: 94) This aspect is criticized in Collier (2000), who substitutes a dynamic concept of “cohesion” for autopoiesis because it admits of degrees and does not sacrifice interaction with the environment in the explanation of autonomy.

²³ Swenson (1992: 209) Cf., *ibid.*, on the “tautological” character of explanation in the theory of autopoiesis.

both explicitly depart from the methods of evolutionary biology under the guiding idea that they cannot adequately address the question of what “life itself” is, and both propose a type of metabolic closure as an answer. Like autopoiesis, the theory of (M, R) systems approaches the problem starting from a strong opposition between relational and dynamical analysis (though this opposition is much more rigorously developed in Rosen’s work). Beyond their general similarities, however, the two theories contrast in ways which are highly instructive for our guiding theme. The theory of (M, R) systems builds on itself throughout the intricately interconnected sections of *Life Itself*, and would be impossible to fully present here. We will accordingly be limited to presenting only a brief summary, limited to the hylomorphic features of Rosen’s theory.

Before approaching the central argument, it will be necessary to introduce some of Rosen’s distinctive ideas on science and mathematical modeling. As Rosen argues, science works within language and according to the duality in language between syntax—propositions and their production rules internal to the language—and semantics, the meaning and reference of linguistic units. Constructing a purely syntactical language is desirable, as Rosen writes, because it can enable the “inferential entailment” of propositions from each other using production rules, without requiring one to consult the meaning or reference of the propositions at all (1991: 43). As is well-known, however, a natural language can never be completely formalized, or reduced to syntax. There can nevertheless be “sublanguages” within a natural language which are purely syntactical. Rosen understands the operation of theoretical science to be analogous to the extraction of a syntactical sublanguage from a natural language: science extracts and builds “formalisms,” or mathematical, syntactical languages, from the natural world (1991: 43-44).

Rosen’s central argument proceeds from the idea that the formalism of Newtonian dynamics is so restrictive and so “poor in entailment” that it will never allow us to grasp what life is (1991: 99). Yet, in his estimation, working within this formalism is still virtually synonymous with what it means to do science. The formalism of Newtonian dynamics in Rosen’s view responds to the following problem: given a set of physical variables, it is possible to observe correlations in the changes among their values, but not to entail the values necessarily from other values. That is, given a formal proposition, what is sought are the “production rules”—the laws of motion—that entail one proposition from another. This is accomplished in two steps: (1) restricting the proposition to just two variables (position and velocity), which define the instantaneous state of the system; and (2) defining the production

rule as a function of that state, or force as a function of acceleration.²⁴ The result is a “recursive entailment of states,” where each state is necessarily entailed by the preceding state. The strict necessity won through this operation imposes severe restrictions, however: anything which cannot be encoded into the instantaneous state or the dynamical law cannot figure into the causality determining changes in the system.

Rosen contends that dynamical methods may be able to describe the partial mechanisms within life, but cannot grasp the living system as a whole. In order for the latter to be possible, a new kind of formalism—a mathematics not based on recursion rules and a physics not based on dynamics—is necessary. The mathematics Rosen uses for this purpose is elementary category theory, a highly abstract formalization of mathematical language similar to set theory. Category theory is a compact language that may be used to describe sets, containing members or elements, and mappings, or transformations from one set to another.²⁵ Such mappings are called, suggestively, “components,”²⁶ in that they take an input, an element from a set, and transform it into another element of another set as an output. This is written as $f: A \rightarrow B$, where f is the mapping or component and A and B are the sets containing elements a_1, a_2, \dots , and b_1, b_2, \dots . A group of components and the elements they transform may be graphed in block diagrams consisting of components, inputs and outputs (common in cybernetics and information sciences). Now, this formalism does not yield any recursive entailment, but for this reason it is much less restrictive and allows other kinds of entailment to be modeled. Rosen interprets the different entailments involved in the simple system $f: A \rightarrow B$ in terms of Aristotelian causality. Given that the component f transforms a into b , one can regard f as the efficient cause of b , and a as the material cause of b . Final causality may also be present in the form of “functions” that f takes on when it is embedded in a larger system of components.²⁷ Rosen’s category theory with its modes of entailment is thus supposed to provide a language for modeling the organization of systems. Such organized systems are distinguished by their relational and functional character—i.e., whereas every dynamical system is absolutely unique and particular, organized systems produce certain general kinds of effects which are, to some degree, independent of the particular material

²⁴ These two steps correspond to Newton’s first and second law, respectively. The third law allows for the composition of multiparticle systems from the rules governing single-particle systems.

²⁵ Of course, category theory is more complex than this, and Rosen’s treatment of category theory in *Life Itself* (1991) is arguably somewhat basic. Readers interested in looking into category theory in greater depth are referred to textbooks on the subject: Lawvere and Schanuel (1997) and Pierce (1991).

²⁶ This notion of “components” is distinct from “components” in the autopoietic theory, which were the same as “parts.” Here, they are to be understood like the autopoietic “processes of production,” or more generally as “relations” among parts.

²⁷ The formal cause is not given much explicit treatment, but it would be represented by the whole group of mappings like f in which b is embedded (if there is such a group).

realization of the system. Hence, the types of systems amenable to Rosen's relational formalism would preeminently be functionally organized systems: human-made machines and organisms.

This relational formalism is exactly what Rosen opposes to the formalism of dynamics. The strategy of the latter is reductionistic: approached with a natural system, it aims to "throw away the organization and keep the underlying matter" (1991: 119). It is then hoped that, after enough work is done, the organization can eventually be recaptured or reduced. The strategy advocated by Rosen, on the contrary, is "relational": it aims to "throw away the matter and keep the underlying organization" (119).²⁸ Each approach, being a formalism, loses referents in passing from the natural world to the model—losing organization in the one case, and matter in the other case. Unlike reductionism, however, the relational approach does not attempt to recapture the referent by adding enough "syntax" (i.e., by complete formalization). Instead, it is hoped that it can recapture the matter through a process of "realization," where a matter is found to realize the form (120). This is essentially a variation on the distinction between a "bottom-up" approach and a "top-down" approach as expressed above.

So far the opposition between relational and dynamical analysis closely parallels the expression of this opposition within the theory of autopoiesis. The essential difference, however, is that in Rosen's theory, applying the relational analysis does not confirm the cybernetic postulate that organisms are types of machines. Let us look again at the simple machine $f: A \rightarrow B$, consisting of a single component transforming inputs into outputs. The machine's output or effect b is causally explained or "entailed" by the structure of the machine, but the transformation f is not. It must be explained by reference to the machine's environment. If $f: A \rightarrow B$ is an organized system, however, this will involve an inexplicable jump unless we would also ascribe some sort of organization or finality to the environment. For this reason, Rosen argues, so long as the efficient cause of an organized system lies outside of it, it must lie in some functional or purposive context of fabrication (1991: 246).

For organisms rather than machines, however, f may be entailed from within the system without leaving that new component itself unexplained and occasioning an infinite regress. Instead of entailing f by adding components to the system, one adds more "relations of entailment," or organization, to the components already present. Rosen here makes use of the peculiar reflexivity of category theory notation, whereby, for instance, mappings may have as their input or output other sets of mappings, and elements of sets may select mappings from sets of mappings (1991: 134-8). It should be possible in principle, then, to exhibit the

²⁸ Rosen also writes, "throwing away the physics and keeping the underlying organization" (1991: 280).

components of a system as causing each other without leading to an infinite regress. This would logically require that at least one component would act as efficient cause of two other components simultaneously (Cárdenas et al. 2010: 86), or that at least one element would act simultaneously as efficient cause and material cause. This cannot be the case in a machine, where causality is always “fractionable” into distinct parts (Rosen 1991: 212). For Rosen, however, it is the organism’s most fundamental property.

The minimal living system, the (M, R) system, is constructed in this way as a circuit of overlapping sets and mappings in which each component is entailed:

$$f: A \rightarrow B; \quad \Phi: b \rightarrow f; \quad \beta: f \rightarrow \Phi.$$

Here f represents metabolism, Φ repairs f using an output of metabolism, and β , a “property” or “molecular configuration” in B , replicates the entire system (Rosen 1972).²⁹ Together, due to their mutual entailment, these components form a system that is “closed to efficient causation” (Rosen 1991: 244). That is, all of the catalysts governing reaction pathways, all of the topological properties determining the system’s organization, are generated from within the system itself.

This result would therefore define life by metabolic closure in a way that is similar to autopoiesis but more rigorously deduced. It is particularly interesting how the formal language of category theory reveals the combinatorial requirements of such a system: namely, that closure in a system of parts and their relations, or sets and mappings, requires going beyond this very duality at at least some points. It thus becomes evident how the central result of Rosen’s theory of (M, R) systems overcomes what was the primary deficiency of the autopoietic theory: namely, its separation between the material parts of the system and its relational organization, which lay behind its recourse to formal causality and explanations by definition. In providing a formal language for the autopoietic “processes of production of components regenerating the processes...,” Rosen’s theory gives a more definite sense to the reciprocal causation between the parts and the relational form of the system. Moreover, it is shown that this aspect decisively distinguishes the living system from the machine. In Rosen’s (M, R) system, the “functions” which components take on, in relation to the other parts of the system, are generated and “selected” from the bottom up. In a machine, however, the functions inhering in its relational organization are specified from the outside, in a purposive context. This kind of determination of the form of organization from the top down, and the

²⁹ Here, as per Rosen’s suggestion, a component such as f may be regarded as a set of catalysts facilitating the transformation of chemical sets $A \rightarrow B$. Note that Rosen’s notion of repair has nothing to do with DNA repair, and it has been suggested to call this “replacement” instead in Letelier et al. (2006: 950). The notion of replication also differs from how that word is used today in connection with DNA replication.

concomitant separation of form and matter, would appear in light of Rosen's central result to be characteristic of the fabrication of machines rather than of the generation of organisms.

In spite of this important conclusion, we must bear in mind that Rosen's central result defines a purely formal system, and that the elements B, f, Φ , etc. are formal objects having "syntactical" coherence but no "reference." With this formal (M, R) system, we therefore have a different kind of hylomorphism than that of the autopoietic theory. Instead of a division between relations and parts within the domain of "syntax," to use Rosen's terminology, there is a division between the whole formal system of parts and relations, on the one hand, and its realization, on the other, or between the "syntax" and its "reference."³⁰ Between these terms there is, moreover, a secondary, extrinsic, or "arbitrary" relation—a *rapport* in Simondon's sense—because the particularities of the realization are not "entailed relationally" (1991: 140). There is no mapping between the formalism and its realization (1991: 280), just as the blueprint of a machine cannot give a complete set of instructions for its construction.

The strategy of Rosen's theory of (M, R) systems strongly embodies the top-down approach described above. As we noted, the top-down approach depends on the conditions that (1) a realization can actually be found for the formalism, and that (2) the features of this material realization will not have to enter into the determination of the formal relations defining the system. Regarding the first condition, it is evident that the formalism constructed by Rosen does not allow us to say very much about a real system of metabolic closure. Realization cannot be a matter of directly choosing real chemical sets corresponding to B, f, Φ , etc. (Rosen 1972: 251.) It may therefore be a pragmatically better strategy of realization to start "from the bottom up" by looking at examples of metabolisms and then abstracting the general properties of metabolic closure.³¹ Whether a realization can be found for the (M, R) system is a question that probably can only be settled *a posteriori* (Cottam et al. 2007: 2364). Thus far, the theory of (M, R) systems has not seen much appreciable success with respect to the realization problem, due most likely to its high degree of mathematical abstraction.³²

³⁰ This distinction between syntax and reference, or formalism and realization, is otherwise expressed in a later essay of Rosen's in *Essays on Life Itself* (2000) in terms of the philosophical distinction between essence and existence. Whereas in science one usually starts from the existence of some natural phenomenon and tries to determine its properties (essence), with (M, R) systems we start from the essence, or the formalism, and try to either show how it exists or make it exist (256-264). This unequivocally underscores Rosen's implicit hylomorphism, in which the essence of the individual consists in its form.

³¹ This is, to some extent, an approach exhibited in Morowitz and Smith (2007), and also Gánti (1997).

³² J. C. Letelier et al. (2003) write: "Neither Autopoietic systems nor (M, R) systems have been used to explain any experimental findings or to predict new biological phenomena in an unambiguous way. It is not surprising then that these theoretical models have been neglected by the vast majority of experimental biologists. This neglect may reflect the fact that both theories are incomplete in the fundamental aspect of how to map their theoretical concepts (structure, organization, Φ , circularity, etc.) with experimental entities." (267) A similar reservation is expressed in Soto-Andrade et al. (2011).

Regarding the second condition, it is Rosen's contention that the form of the (M, R) system can be determined while "throwing away the matter," or independently of the matter. As we noted, however, the form of the (M, R) system—closure to efficient causation—belongs to a system of formal objects without referents. The independence of the form may therefore not hold under conditions of *metabolic* as opposed to category closure. The primary reason for supposing this is that the elements of Rosen's notation are indifferently connectible combinatorial objects, which can thus be defined as forming the organization one likes independently of any material specificities. The independence of the form is in this sense merely a consequence of the formalism of the notation itself. The elements of metabolism—chemical compounds composing metabolites, enzymes, etc.—have, on the contrary, limited chemical affinities, which specify "from below" the types of relations they may enter into with other compounds. A variety of chemical and thermodynamic properties inhere in the matter which determine the form of organization the matter may produce. Of course, using Rosen's terms, the part does not strictly "entail" its functional place within the whole system, but neither is there an "arbitrary" relation between the two (1991: 140). This is in fact what is suggested by the extreme degree of entailment among the components and material inputs and outputs of the (M, R) system.

It would be fair to conclude on the basis of the preceding that the theory of (M, R) systems is hylomorphic in the sense we have described. While there is little by way of formal causality in the theory, there is certainly a strong separation between the form and the matter, conceived here as formalism and material realization. Accordingly, as in the theory of autopoiesis, the form is conceived of in a static rather than generative manner: there is an external or "unentailed" relation (*rapport*) between the form and the matter, and no intermediary order between the relational form and the material parts. The hylomorphism of the theory seems to ultimately derive from Rosen's views on scientific modeling and entailment. Science for Rosen extracts formal languages, or *mathematical* systems of entailment, from the natural world; and without this kind of strict "syntax" of entailment we have mere metaphor instead of models (1991: 64-66). In effect, this opens a wide gap between the allowable formalisms possessing a strict syntactic entailment, i.e., between the recursive entailment of the Newtonian formalism and the relational entailment of (M, R) systems.³³ This requirement ends up being restrictive on both sides: on the one hand, we have a purely formal notion of life whose reference or realization remains uncertain, and on the other hand, we have a

³³ Some prominent attempts in the 20th century to bridge this gap—Ludwig von Bertalanffy's work on "open systems," Prigogine's nonlinear thermodynamics, and René Thom's "catastrophe theory"—are denounced by Rosen as being metaphorical (1991: 65).

classical-mechanistic notion of matter which is valid only in the domain of closed systems and conservative forces. Inquiring into the intermediary order lying between these extremes will obviously have to involve considerations from chemistry, and it is here that the chemoton theory of Tibor Gánti is instructive.

3.3 THE CHEMOTON

The chemoton, or “chemical automaton,” is the model of minimal life developed by the Hungarian biochemist Tibor Gánti (1971, 1975), and presented in detail in the book *The Principles of Life* (2003). Like autopoiesis and (M, R) systems, Gánti’s chemoton theory approaches the question “what is life?” in terms of a minimal metabolic system. It also implicitly defines the living system by a kind of closure—or rather by the “autocatalyticity”—of chemical reaction cycles. There are nevertheless important differences between chemoton theory and the other models considered so far. The most important for our purposes is that the chemoton is, in Gánti’s own words, “designed from the bottom up, starting from the direction of the chemical reactions” (2003: 5). The chemoton is the most strongly based in chemistry of the three models we consider, and we will see that this has important consequences for how it stands in relation to hylomorphism. There is also more included in the chemoton than in the other models: it comprises a circular metabolism, but also coupled to this, an autocatalytic membrane and an information-template subsystem. Like autopoiesis and (M, R) systems, chemoton theory represents an elaborate body of work, and it will only be possible to present its central features in comparison with the other models, as it is relevant to the guiding idea of hylomorphism.

The basic point of departure for the chemoton theory is the recognition of the particular material constitution of living things—their chemical and fluid nature. Like machines, living systems differ from the physical environment in that their structure directs the flow of energy along constrained paths to perform “useful” work, instead of performing work in a random or undirected way (Gánti 2003: 2). They are both “functional” systems in the sense that they have a reliable manner of “functioning” or operating, producing specific types of effects. Living systems nevertheless differ in a crucial way from the more familiar human-made machines utilizing mechanical and electrical energy. These latter machines consist of hard parts possessing a rigid “geometrical order” (e.g., wheels, gears, springs, wires, chips, etc.) (ibid.). Mechanical or electrical energy is constrained along definite pathways in such machines due to the discrete extension of the parts, or the spatial arrangement of the gears,

springs, pistons, etc. In living systems, by contrast, the driving force is the chemical energy of reactions proceeding in a fluid phase, where the relevant factors are not extensive, but intensive and qualitative—e.g., chemical concentration and affinity. Their parts are not hard and solid, but “soft” and “fluid.” For Gánti, living systems are species of “fluid automata,” like the complex reaction networks used in some industrial chemical processes, or like oscillating chemical cycles such as the well-studied BZ reaction.

For Gánti, then, the living system is a type of machine, but its fluid, chemical material composition differentiates it from the machine as we typically understand it. Gánti’s theory thus utilizes a somewhat idiosyncratic concept of the machine. It is not defined by reference to purposiveness, by its relation to human aims, nor by the realization of a separable design or blueprint: it is rather an “automaton” in the sense of something which operates, or functions, due to constrained flows of energy performing work in a directed way. This is distinguished from the cybernetic notion of the machine present in the theories of autopoiesis and (M, R) systems, where the machine was defined by the separability of a relational, functional organization. For autopoiesis and (M, R) systems, the machine-character of life justifies “throwing away the underlying matter” in favor of the relational organization (even if this leads Rosen to make a different kind of distinction between living things and machines). In Gánti’s chemoton theory, by contrast, the essential determinant of the living thing as a fluid automaton is precisely its matter. Hence, the primary distinction between the living and the non-living is not mapped onto the distinction between form and matter, as it is in the other theories, but begins from the type of “matter” determining a system from the bottom up.

This material difference between fluid and solid automata can, moreover, partially account for some of the distinctive properties of living systems described in the other theories, such as their dynamic homeostasis and their extra relations of “entailment” allowing them to self-repair. The cycling of chemical reactions, for instance, grants fluid automata a degree of homeostatic stability against perturbations. Moreover, since they are “largely independent of any geometrical order,” they may still function even while growing or being divided, making it possible for them to reproduce or self-repair in ways that are completely inaccessible to hard automata. Gánti invokes Leibniz on this distinction between soft and hard, fluid and solid: whereas artificial automata are not functional machines in their smallest detail, natural or divine automata (living bodies) are functional machines up to the limit of the infinitely small (2003: 12-13),³⁴ or in their smallest parts. The “smallest parts” would here be the minimum reaction network—or, at the idealized limit, a single instance corresponding to the

³⁴ Cf., Leibniz (1925), §64.

chemical equation for the network. In this sense, if this instance would itself be a fluid automaton, at least in principle, then one could say that the “form” of the fluid automaton is nothing other than the chemical equation giving the reaction specificities which determine the network. The “form” of the fluid automaton would already inhere in this minimum of parts. Stated in another way, the channeling or constraint of energy is not arranged as in hard automata from above, by the discrete arrangement of the parts with their rigid boundaries. Instead, constraint is distributed throughout the fluid system, and the flow of energy is directed along the direction arrow of chemical reactions. The form of the fluid automaton, its relational organization, is thus only the set of relations internal to its matter.

The fluid and chemical nature of the chemoton model bears a fundamental philosophical significance. As James Griesemer notes in his preface to *The Principles of Life*, the chemoton “introduc[es] a seriously chemical way of thought into the philosophy of biology” (viii), as opposed to, e.g., a relational way of thought, or the more common “informational” way of thought. But what would it mean to “introduce a chemical way of thought,” if not simply a call for reductionism? This is how Rosen characterizes modeling strategies that focus on matter rather than organization (1991: 119). In fact, the chemical nature of Gánti’s fluid automata does lead to a sort of “reduction,” but it is not a reduction in Rosen’s sense of a formalized, recursively entailed series of states. This is because Gánti does not rely on the methods of chemistry that are strongly based on dynamics—i.e., chemical kinetics, where reaction rates and chemical concentrations play an analogous role to acceleration and state, respectively. If he did, we would indeed have a kind of recursive entailment of states as Rosen describes it.³⁵ Instead, Gánti utilizes stoichiometric equations, which give the relative proportions of reactants and products found in the chemical equation, modified by Gánti to handle the nonlinearity of chemical cycles. In Gánti’s innovative “cycle stoichiometry,” as he writes, “only the quality and quantity of the chemical components and the pathways of their formation, transformation and decomposition are described” (2003: 5). As distinguished from the dynamical formalism, then, it is basically indifferent to the concepts of state and dynamical law, concentration and reaction rate (2003: 5-6). Hence, it cannot be reductive in the sense that mechanics is reductive.

It could be characterized as being reductive in another sense, however. As Griesemer writes, Gánti’s stoichiometric approach has the effect of advancing “flow of matter as a fundamental property of biological systems” (2003: 177), counseling us to ““follow the stuff”

³⁵ The crucial difference would be of course that the evolution of a chemical system would be thermodynamically irreversible and stochastic, whereas that of a purely mechanical system would be reversible and deterministic.

as it moves around the chemical network” (185), irrespective of the organizational relations that are supposed to obtain in the system. All of the parts of the chemoton fall under the discipline of having to satisfy the stoichiometric mass balance equations, which describe the direction of flow of matter through the chemoton. In this manner, Gánti’s chemical and stoichiometric approach poses the requirement that the abstract expression of all of the system’s higher-level features, including its organizational or relational properties, must be fulfilled in terms of the flows of matter through the chemical cycles described in the model. Perhaps, then, it could be called “reductive” in the etymological sense of “leading back” the higher achievements of the fluid automaton to a monistic layer of chemical interactions. As an example, the idea of genetic information providing “instructions” to the other elements of the cell would be something to be “reduced” in this way.³⁶ Crucially, however, this kind of reduction is not reductionistic, at least not in any traditional sense: it is not argued that biology is reduced to general chemical laws, but only that this chemoton model is based on an order of chemical interactions under the highly specific conditions. This becomes especially evident when we consider the specific organization of chemical cycles in the distinct subsystems of the chemoton.

The chemoton is not just a fluid automaton, like the BZ reaction or any other chemical dissipative structure. So far we have described only the fluid automaton and cycle stoichiometry aspects, but they constitute only a small part of the theory. The chemoton is a fluid automaton having a highly specific organization. It is composed of three distinct fluid automata, which are each autocatalytic cycles coupled to each other stoichiometrically. The first and most fundamental of these is the metabolic subsystem, which is the non-equilibrium chemical “motor” of the chemoton, providing material to the other subsystems and maintaining the homeostatic stability of the whole. Coupled to this is an information subsystem which produces larger template polymers using materials from the metabolism. These template molecules can regulate the amount of material produced in the other subsystems, and can also carry hereditary information. Finally, a membrane system is coupled to both of these which encloses the informational and metabolic cycles within their suitable range of chemical interaction. The chemoton is significantly different from even the simplest modern cells. Perhaps the most important difference would be that regulative control in the chemoton is distributed in the stoichiometric connections among the subsystems, instead of being more localized in the specification of proteins from the DNA sequence. It is nonetheless not a purely abstract or formal model. The basic “functions” of modern cells are contained in

³⁶ Cf., Griesemer (2003: 185).

the chemoton in a more primitive way, but instead of remaining abstract or formal terms their meaning can be fulfilled in terms of underlying chemical interactions pictured in the model.

One of the most generally useful results of chemoton theory is that the three-subsystem model allows Gánti to develop a clear set of “life criteria” for answering the question “what is life?” Gánti distinguishes real life criteria, criteria for the state of “living,” from a set of potential life criteria, criteria for the existence of a living world. The chemoton is a bridge between these two sets of criteria. His real life criteria include that the living system is a unified whole, performs metabolism to maintain non-equilibrium, is inherently stable, contains information about itself, and regulates and controls its own processes. Among the potential life criteria are included that the living system is capable of growth and reproduction, capable of hereditary change and evolution, and mortal. These life criteria enable Gánti to deal with the limit cases of life in a clear manner. Paradoxically, however, they also show that life is marked by a variety of properties that only partially overlap, and that the selection of essential properties is, to a significant degree, relative to the stage considered within an ongoing process of transitions.³⁷ The chemoton is a unit of early evolutionary transitions rather than the definitive essence of life.³⁸

Chemoton theory is especially valuable for being able to combine conceptual depth on the fundamental question “what is life?” with a bottom-up approach grounded in the dynamics of chemical cycles. In fact, these two aspects inform each other in a subtle way. With Gánti’s bottom-up approach, one does not start from a flat notion of matter, which then necessitates locating all of the important “lifelike” properties in the relational organization of the system. Instead, starting from the norm of matter in a more “lifelike” state, we see an intermediary order between matter as described by mechanics and life fully organized. This would be the intermediary order of fluid automata, chemical dissipative structures, and the pre-biotic chemical cycles described elsewhere by Gánti (1997: 590) and by such authors as Morowitz and Smith (2007: 53). Within this perspective, there are constitutive properties of life in addition to its having the form of metabolic closure. Approaching the problem in this way in effect narrows and displaces the distinction between living and non-living nature, and precludes it from becoming a distinction wholly centered around the possession of a given relational form as it is in autopoiesis and (M, R) systems.

³⁷ As Szathmáry points out in his commentary, “One could argue that [a system with only the metabolism and membrane subsystems] could be regarded as an ‘even more minimal’ model for life than the chemoton” (2003: 162). If the membrane function were provided by abiotic features of the environment, moreover, we would have a purely metabolic system, which it is difficult to decide whether it would be living or not.

³⁸ Cf., Griesemer (2003: 173-5)

4 DISCUSSION AND CONCLUSION

Fundamental theories of life, such as the three we have examined, are theories that attempt to describe the most fundamental properties that ground and explain the differences between living and non-living systems. Such theories can be construed as aiming to characterize the definition or the “essence” of life, or to distinguish between essential and accidental amid the number of properties described by biologists.³⁹ The difficulties surrounding this enterprise are manifold and well-known.⁴⁰ Many of these difficulties revolve around the fact that we are looking for a universal definition of life that is not limited to life as we know it today, even though it is not known how far life as such can possibly differ from the actual life we know. Under these conditions, a conservative strategy would be to stick to the actual life we know, where definitions can be falsified. This roughly corresponds to the descriptive approach to arriving at philosophical definitions, which attempts to abstract a general definition from empirically known instances of the *definiendum*. While adequate for some purposes, it is not sufficient for defining the essence of life, for reasons of principle and of fact. First, in principle, the descriptive approach relies on our intuitive recognition of instances, but there is no guarantee that our intuitions form a coherent whole. There are divergent instances corresponding to intuitive notions of life. The definitions that are abstracted from them will therefore only partially coincide, and the general definition will have to accommodate this difference by either becoming excessively inclusive or arbitrarily exclusive. This explains why the limit-cases (viruses, non-reproducing organisms, whole ecosystems, etc.) tend to confound attempts to reach a descriptive definition of life. Second, in fact, there is a broad consensus that early life on Earth must have been considerably different and simpler than life as it exists today. What is actual today may not resemble what was actual in the deep past, nor what might have been actual. These particular problems may be overcome or at least

³⁹ The notion of “essence” here should not be rejected out of hand on the basis of non-necessary metaphysical connotations. In order for there to be an essence of life, it is enough that there should be at least one property common to all living things but absent from non-living things. Whether there is such an essence is a matter of how well-defined in nature the division is between life and non-life, and need not involve any non-naturalistic or otherworldly assumptions. Of the different accounts of essential properties, our use of the notion is closest to the explanatory account (cf. Gorman 2005, Copi 1954), which is arguably the most appropriate for natural entities such as living systems, rather than the definitional account or the more common modal account. According to this explanatory account, the essential properties of an entity are those that are most fundamental, such that they ground and explain its other distinctive properties.

⁴⁰ Cf., Morange (2008), Luisi (1998), Bedau (1996), etc. for a useful classification and critical assessment of different definitions.

attenuated through a constructive rather than descriptive approach to defining life.⁴¹ Instead of starting from the intuitive recognition of instances, it attempts to construct a minimal instance from simple elements and basic principles known to govern living processes (e.g., chemical principles, functions of homeostasis or repair, etc.). This approach has the advantage of minimizing the instance under consideration in order to reach a compact and unambiguous distinction between essential and accidental properties of life. The disadvantage is evident, however: in departing from empirically known instances of life, the model that is constructed will tend to be highly formalistic and abstract. It is as if the formalism of the model is a cost of its being able to circumvent our lack of knowledge about the exact material conditions of early or minimal life.

The different theories we have considered each define life by some kind of metabolic closure, but they reach this definition in different ways. They can be roughly categorized along two different axes: descriptive versus constructive, and bottom-up versus top-down. The theory of autopoiesis may be said to utilize a more descriptive and top-down approach, which abstracts a general concept from the functioning of cells. While it is useful as a conceptual tool, autopoiesis remains at an abstract level of description which is not fulfilled in terms of concrete material interactions in the model. At the other extreme, the theory of (M , R) systems belongs to a constructive, top-down approach. It is so far removed from the descriptive domain of actual biological entities, however, that the model remains a formal object without reference, and we are consequently left with the formidable “realization problem.” Chemoton theory utilizes a bottom-up approach that is both descriptive and constructive: the chemoton with its subsystems is a constructed model, but it a construction based on examples of chemical cycles and known chemical principles.⁴² For this reason, it is capable of overcoming the problems associated with a purely descriptive approach without remaining a merely formal or abstract model.

I have argued that the formalism of the first two of these models is involved in a hylomorphist background ontology. This ontology has had a long history of influence on biological thought ever since the synthesis of the two in Aristotle’s philosophy. In this respect it is fitting that Simondon characterizes hylomorphism in general as a “vital” model of the individual (2005: 24). But he also characterizes it as a “technological” model—and these two aspects are closely related. In any technology of fabrication there exists a duality between the form or design and its material realization. For machines in particular, it is thought that the

⁴¹ cf. Griesemer (2003: 180-181).

⁴² Cf., Gánti (1997).

form or the relational organization is the essential thing, and that the matter need only be compatible with it. In fact, the idea that the form is essential and the matter is accidental is precisely what justified the analogy between living systems and machines in the theory of autopoiesis, and even to some extent in (M, R) systems. As we have seen, however, the assimilation of life to the machine model is problematic for a number of reasons. There is a crucial difference between the ways form is produced in the two cases: whereas form in nature is the end result of generative processes, in human artifice the form comes first and is followed by its material realization. The latter is, in this sense, a “genesis in reverse.”⁴³

Now, the temporal order of terms would be unimportant if the essence of the individual remained the same no matter how its form is produced. It has been one of the primary aims of this paper, however, to show that this is not the case. Even assuming that the *form* would be the same before and after its realization—which has been shown to be questionable—the order of terms fundamentally influences how the *essence* of life gets determined. Starting from the form and moving “top-down” toward its realization, one arrives at what we have called, using Simondon’s terminology, a “static” rather than or “generative” notion of form.⁴⁴ A static notion of form would be one in which the form is what it is prior to its realization, or one to which a matter can be added without altering its nature. This requirement of a static form was stated above for the top-down approach: namely, that the matter which is found as a realization will not have to enter into the determination of the form; or, that there is an external relation (*rapport*) between the matter and the form. In a generative notion of form, by contrast, the form is just the set of relations internal to the matter in the course of its changes. The form is represented as the contingent result of a process that might have produced something else.

In the fabrication of machines, the identity of the form before and after its realization is secured by the purposive work of the human technician. In the genesis of living things, by contrast, the form cannot be detached from the process and put at the start as its guiding principle without introducing teleology.⁴⁵ Nor can the form, considered as a contingent result, be identified with the essence of the individual. With a generative notion of form, the essence of the individual (if there is such an essence) would have to include *more than* its finished form, because the latter is not determinable independently of the generative processes that produce it. This non-independence gets expressed in Simondon’s generalized model of

⁴³ Cf., Simondon (2005: 23).

⁴⁴ Cf., Simondon (2005: 363).

⁴⁵ In fact, according to one interpretation, this idea lies at the root of Aristotle’s conception of final causality. (Gottlieb 1976) Cf. also McCall (2012: 183-184).

individuation via the conjunction of three necessary individuating factors—a “pre-individual,” metastable source of potential energy that interacts with an “informed” matter to produce the eventual, formed individual. Whatever the specific model of individuation it may be, with a generative notion of form, understanding the form requires placing it within such a generative field, where it loses its independence and autonomy and becomes one potential result among many. The “essence” of the individual becomes of necessity partially ambiguous because it includes more than what belongs properly to “itself,” and certainly more than its form. In brief, with a generative notion of form, as distinguished from a static notion of form, the essence of life cannot consist solely in its form.

The difference between these notions of form corresponds to the contrast we have elaborated between autopoiesis and (M, R) systems, on the one hand, and the chemoton, on the other hand. The formalism of the first two models allows for a clear-cut and unambiguous definition of the essence of life: namely, “autopoiesis” or “closure to efficient causation.” By contrast, the essence of life is more ambiguous and complex in chemoton theory precisely because it is not limited to the form. In fact, I would claim that there is ultimately no clear-cut “essence” of life in chemoton theory. Gánti’s theory presents the form of life within the generative field of fluid automata and the informed matter of pre-biotic chemical cycles, such that the form can be “reduced” or “led back” to the diverse potentials latent in this underlying layer of chemical interactions. The situation is similar to that described by Morowitz and Smith (2007: 51): “A part of the order we recognize as living is thermodynamic order inherent in the geosphere.” The chemoton straddles our conceptual division between life and matter, as it were: it is constituted by more than what properly belongs to itself. From this perspective, it is perfectly natural that strictly formal models will lead to a one-sided emphasis on the biological properties of autonomy and closure rather than those of energy transfer and openness. With a generative notion of form, however, it should be possible to “reduce” these formal figures of life, and thereby to reach a better—albeit more complex—answer to the question “what is life?” than is available in a hylomorphic scheme.

Acknowledgements: This work was made possible by a grant from the Scientific Research Fund of Flanders (Fonds Wetenschappelijk Onderzoek - FWO).

References

- Aristotle (1984) *The Complete Works of Aristotle: The Revised Oxford Translation*, Barnes J (rev). Princeton, Princeton
- Atlan, H, Bousquet, C (1994) *Questions de Vie*. Seuil, Paris
- Barnes, G (2003) The Paradoxes of Hylomorphism. *Rev Metaphys* 56:501-523
- Bedau, M (1996) The Nature of Life. In: Boden M (ed) *The Philosophy of Artificial Life*. Oxford, Oxford, pp 332-360
- Cárdenas, ML, Letelier, JC, Cornish-Bowden, A, Gutierrez, C, Soto-Andrade, J (2010) Closure to Efficient Causation, Computability, and Artificial Life. *J Theor Biol* 263, 79-92. doi: 10.1016/j.jtbi.2011.01.030
- Chandler, J, Van de Vijver, G (eds.) (2000) *Closure: Emergent Organizations and Their Dynamics*. New York Academy of Sciences, New York
- Chyba, CF, Mcdonald, GD (1995). The origins of life in the solar system: Current issues. *Annu Rev Earth Planet Sci* 23:215–249. doi: 10.1146/annurev.earth.23.1.215
- Cleland, CE, Chyba, CF (2002). Defining life. *Origins Life Evol B* 32:387–393
- Collier, J (2000) Autonomy and Process Closure as the Basis for Functionality. In: Chandler J, Van de Vijver G (eds) *Closure: Emergent Organizations and their Dynamics*. New York Academy of Sciences: New York. doi: 10.1111/j.1749-6632.2000.tb06287.x
- Copi, I (1954) Essence and Accident, *J Philos* 51:706–719. doi: 10.2307/2021504
- Cornish-Bowden, A, Cárdenas, M, Letelier, JC, Soto-Andrade, J (2007) Beyond Reductionism: Metabolic Circularity as a Guiding Vision for a Real Biology of Systems. *Proteomics* 7:839–845. doi: 10.1002/pmic.200600431
- Cottam, R, Ranson, W, Vounckx, R (2007) Re-mapping Robert Rosen's (*M,R*)-Systems. *Chem & Biodivers* 4:2352-2368. doi: 10.1002/cbdv.200790192
- Dennett, D (2011) Shall We Tango? No, but Thanks for Asking. *J Conscious Stud* 18:23-34
- Dennett, D (2013) Aching Voids and Making Voids, A Review of *Incomplete Nature: How Mind Emerged from Matter* by Terrence W. Deacon. *Q Rev Biol* 88:321-324. doi: 10.1086/673760
- Eigen, M, Schuster, P (1977) The hypercycle: a principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* 11:541-565. doi: 10.1007/BF00450633
- Etcheberry, A (2004). Autopoiesis and natural drift: Genetic information, reproduction, and evolution revisited. *Artificial Life*, 10, 347–360. doi: 10.1162/1064546041255575
- Gánti, T (1971) *Az élet princípuma*. Gondolat, Budapest
- Gánti, T (1975) Organization of Chemical Reactions into Dividing and Metabolizing Units: The Chemotons. *Biosystems* 7:15-21. doi: 10.1016/0303-2647(75)90038-6
- Gánti, T (1997) Biogenesis Itself. *J Theor Biol* 187:583-593. doi: 10.1006/jtbi.1996.0391
- Gánti, T (2003) *The Principles of Life*. Oxford, Oxford
- Gorman, M (2005) The Essential and the Accidental, *Ratio* 18:276–289. doi: 10.1111/j.1467-9329.2005.00290.x
- Gotthelf, A (1976) Aristotle's Conception of Final Causality. *Rev Metaphys* 30:226-254
- Griesemer, J (2003) The Philosophical Significance of Gánti's Work. In: Griesemer J and Szathmáry E (eds) *The Principles of Life*. Oxford, Oxford, pp 169-186
- Hordijk, W, Steel, M (2004) Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *J Theor Biol* 227:451-461. doi: 10.1016/j.jtbi.2003.11.020
- Jacob, F (1970) *La Logique du Vivant*. Gallimard, Paris
- Johnston, M (2006) Hylomorphism. *J Philos* 103:652-98
- Kauffman, SA (1993) *The Origins of Order*. Oxford, New York
- Lawvere, F, Schanuel, S (1997). *Conceptual Mathematics*. Cambridge University Press, Cambridge, UK
- Leibniz, GW (1925) *The Monadology*, 2nd ed. Latta R (ed and trans). Oxford, Oxford (First published 1683)
- Letelier, JC, Cárdenas, ML, Cornish-Bowden, A (2011) From *L'Homme Machine* to metabolic closure: Steps towards understanding life. *J Theor Biol* 286:100-113. doi: 10.1016/j.jtbi.2011.06.033
- Letelier, JC, Marín, G, Mpodozis, J (2003) Autopoietic and (M,R) systems. *J Theor Biol* 222:261-272. doi: 10.1016/S0022-5193(03)00034-1

- Letelier, JC, Soto-Andrade, J, Guíñez Abarzúa, J, Cornish-Bowden, A, Cárdenas, ML, (2006) Organizational Invariance and Metabolic Closure: Analysis in Terms of (M, R) Systems. *J Theor Biol* 238:949-961. doi: 10.1016/j.jtbi.2005.07.007
- Luisi, P (1998) About Various Definitions of Life. *Origins Life Evol B* 28:613-622. doi: 10.1023/A:1006517315105
- Machery, E (2012) Why I stopped worrying about the definition of life... And why you should as well. *Synthese* 185:145-164. doi: 10.1007/s11229-011-9880-1
- Maturana, H (1980) Autopoiesis: reproduction, heredity and evolution. In: Zelený M (ed) *Autopoiesis, Dissipative Structures, and Spontaneous Social Orders*. Westview, Boulder, CO, pp 45-79
- Maturana, H, Varela, F (1980) *Autopoiesis and Cognition: The Realization of The Living*. Maturana H and Varela F (eds) *Boston Studies in the Philosophy of Science* 42. D. Reidel, Dordrecht
- Maynard Smith, J, Szathmáry, E (1995) *The Major Transitions in Evolution*. Oxford, Oxford.
- McCall, S (2012) The Origin of Life and the Definition of Life, in *Contemporary Aristotelian Metaphysics*, ed. Tahko, T. Cambridge, New York, pp 176-186
- Morange, M (2008) *Life Explained*. Cobb, M and Debevoise, M (trans). Yale, New Haven
- Morowitz, H and Smith, E (2007) *Energy Flow and the Organization of Life*. Wiley Interscience 13, 1: 8 October. doi: 10.1002/cplx.20191
- Mossio, M, Moreno, A (2010) Organisational Closure in Biological Organisms, *Hist Phil Life Sci* 32:269-288
- Mossio, M, Longo, G, Stewart, J (2009). A computable expression of closure to efficient causation, *Journal of Theoretical Biology*, 257, 3, 489-498. doi: 10.1016/j.jtbi.2008.12.012
- Nicholson, D (2013) *Organisms ≠ Machines*. *Studies in History and Philosophy of Biological and Biomedical Sciences* 44:669-678
- Oyama, S (2000) *The Ontogeny of Information*. Duke, Durham
- Pierce, B (1991) *Basic Category Theory for Computer Scientists*. MIT Press: Cambridge, MA
- Prigogine, I, Nicolis, G (1977) *Self-Organization in Nonequilibrium Systems*. Wiley, New York
- Rea, M (2011) Hylomorphism Reconditioned. *Philosophical Perspectives* 25:341-358. doi: 10.1111/j.1520-8583.2011.00219.x
- Rosen, R (1972) Some relational cell models: the metabolism-repair system. In: Rosen, R (ed), *Foundations of Mathematical Biology*. Academic Press, New York
- Rosen, R (1991) *Life Itself*. Columbia, New York
- Rosen, R (2000) *Essays on Life Itself*. Columbia, New York
- Ruiz-Mirazo, K, Moreno, A (2012). Autonomy in evolution: from minimal to complex life. *Synthese*, 185(1), 21–52. doi:10.1007/s11229-011-9874-z
- Simondon, G (1995) *L'individu et sa genèse physico-biologique*. Jérôme Millon, Grenoble
- Simondon, G (2005) *L'individuation à la lumière des notions de forme et d'information*. Jérôme Millon, Grenoble
- Soto-Andrade, J, Jaramillo, S, Gutiérrez, C, Letelier, JC (2011) Ouroboros Avatars: A Mathematical Exploration of Self-Reference and Metabolic Closure. In: Lenaerts, T, Giacobini, M, Bersini, H, Bourguine, P, Dorigo, M, and Doursat, R (eds) *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. The MIT Press, Cambridge MA, pp 763-770
- Swenson, R (1992) Autocatakinetics: Yes—Autopoiesis, No: Steps Toward a Unified Theory of Evolutionary Ordering. *Int J General Systems* 21:207-228
- Szathmáry, E (2003) The Biological Significance of Gánti's Work in 1971 and Today. In: *The Principles of Life*. Oxford, Oxford, pp157-168
- Szathmáry, E (2005) In search of the simplest cell. *Nature* 433:469-470, 3 Feb 2005. doi: 10.1038/433469a
- Szostak, J, Artel, D, & Luisi, P (2001). Synthesizing life. *Nature*, 409, 387–390. doi: 10.1038/35053176
- Tahko, T (ed) (2012) *Contemporary Aristotelian Metaphysics*. Cambridge, New York
- Varela, F (2000) *El fenómeno de la vida*. Dolmen Esayo, Santiago
- Varela, F (1994) Autopoiesis and a Biology of Intentionality. In: McMullin B, Murphy N (eds) *Autopoiesis and Perception, Proceedings of a workshop held at Dublin City University on August 25th & 26th, 1992*
- Wiener, N (1961) *Cybernetics: Or Control and Communication in the Animal and the Machine*, 2nd ed. MIT, Cambridge, MA
- Woese, C (2004) A New Biology for a New Century. *Microbiol Mol Biol R* 68:173-186. doi: 10.1128/MMBR.68.2.173-186.2004